

基于节点重要性与相似性的标签传播算法^①



林天森, 孙飞翔

(华南师范大学 计算机学院, 广州 510631)

通讯作者: 林天森, E-mail: lin1210375420@163.com

摘要: 标签传播算法是一种常用的社区发现方法, 具有近似线性的时间复杂度, 但该算法存在随机性和不稳定性. 为了解决标签传播算法存在的准确性低和稳定性差的问题, 本文提出了基于节点重要性与相似性的标签传播算法 (Label Propagation Algorithm based on node Importance and Similarity, LPA_IS). 首先, 基于节点重要性提出种子节点集和算法更新序列的获取方法. 其次, 利用节点重要性与相似性提出了一种计算标签综合影响力的方法, 任意节点根据其邻居标签的综合影响力更新自身的标签. 在真实网络和人工合成网络上进行实验, 结果表明, 与其它 5 种典型标签传播类算法对比, LPA_IS 算法能够在一定程度上提高算法的准确性和稳定性, 并且能够减少算法的迭代次数.

关键词: 社区发现; 标签传播算法; 节点重要性; 相似性

引用格式: 林天森, 孙飞翔. 基于节点重要性与相似性的标签传播算法. 计算机系统应用, 2021, 30(10): 218-223. <http://www.c-s-a.org.cn/1003-3254/8132.html>

Label Propagation Algorithm Based on Node Importance and Similarity

LIN Tian-Sen, SUN Fei-Xiang

(School of Computer Science, South China Normal University, Guangzhou 510631, China)

Abstract: The label propagation algorithm, a commonly used community discovery method, has approximately linear time complexity but randomness and instability. To solve the problems of low accuracy and poor stability of the label propagation algorithm, this study proposes an improved Label Propagation Algorithm based on node Importance and Similarity (LPA_IS). First, based on node importance, a method is proposed to obtain the seed node set and the algorithm update sequence. Second, a method is proposed with node importance and similarity to calculate the comprehensive influence of labels. Any node updates its own label according to the comprehensive influence of its neighbor labels. Experiments on real networks and synthetic networks have shown that compared with five typical label propagation algorithms, LPA_IS can improve the accuracy and stability to a certain extent and reduce the iterations.

Key words: community discovery; Label Propagation Algorithm (LPA); node importance; similarity

在现实世界中, 存在着各种各样的复杂网络, 如: 社交网络、万维网、神经网络等. 这些网络可以用图来表示, 对网络的研究可以转化为对图论问题的研究. 如图论中的聚类问题, 目的是挖掘复杂网络的社区结构, 使得类内节点联系紧密, 类间节点联系稀疏. 挖掘

复杂网络的社区结构对于研究网络结构和应用具有非常重要的意义, 如识别社交网络中的顶级传播者^[1], 根据用户关系进行兴趣或好友推荐^[2,3]和通信网络中的故障恢复^[4]等.

为了挖掘复杂网络的社区结构, 学者们提出了许

① 基金项目: 国家自然科学基金 (61370003)

Foundation item: National Natural Science Foundation of China (61370003)

收稿时间: 2020-12-28; 修改时间: 2021-01-25, 2021-02-07; 采用时间: 2021-02-23

多社区发现的算法. Girvan 等人^[5]提出了 GN 算法, 该算法的主要思想是每次删除边介数最高的边, 直到网络中每个节点表示一个社区为止, GN 算法的时间复杂度为 $O(m^2n)$, 其中, m 表示网络的边数, n 表示网络的节点数. Newman 等人^[6]提出了一种模块度最大化的贪婪算法 FN, 该算法的时间复杂度为 $O((m+n)n)$. Pons 等人^[7]基于随机游走提出的 Walktrap 算法, 其主要思想是采用随机游走算法计算节点的相似性矩阵, 根据相似性矩阵合并社区, 直到所有节点合并为一个社区为止. Raghavan 等人^[8]提出了标签传播算法 (Label Propagation Algorithm, LPA), 该算法的主要优点是时间复杂度低, 适用于在大型网络上的社区发现.

多数典型标签传播类算法在标签更新时, 没有考虑节点重要性和相似性对节点更新的综合影响, 可能导致影响力小的节点反过来影响具有较大影响力的节点, 造成结果准确性低. 本文基于节点重要性与相似性提出了一种改进的标签传播算法 (Label Propagation Algorithm based on node Importance and Similarity, LPA_IS). LPA_IS 算法根据节点重要性提出了种子节点集和算法更新序列的获取方法, 并为任意种子节点分配唯一的标签. 在标签更新过程中, 基于节点重要性与相似性提出了一种标签综合影响力的计算方法, 任意节点根据其邻居标签的综合影响力更新自身的标签, 提高算法的准确性和稳定性.

1 相关工作

LPA 算法的主要思想是, 首先为图中每个节点分配唯一的标签, 标签代表节点所在的社区. 当节点标签更新时, 根据式 (1) 选择邻居标签中数量最多的作为节点的待更新标签, 当邻居中标签数量最多不止一个时, 则随机选择一个. 重复上述操作, 直到所有节点的标签趋于稳定或达到最大迭代次数时, 则停止更新.

$$l_i = \arg \max_{l \in T(i)} \sum_{j \in N(i)} \delta(l, l_j) \quad (1)$$

其中, $T(i)$ 表示节点 i 邻居中标签的集合, $N(i)$ 表示节点 i 的邻居节点集, l_j 表示节点 j 的标签, $\delta(l, l_j)$ 表示克罗内克函数^[9], 如果 $l=l_j$ 时, 则 $\delta(l, l_j)=1$, 否则 $\delta(l, l_j)=0$.

在 LPA 算法中, 初始化序列和标签选择时具有随机性, 导致社区划分时存在准确性低和稳定性差的缺陷. 为了提高 LPA 算法的准确性和稳定性, 近年来, 许多学者对 LPA 算法进行改进, 大致分为以下 3 种. 基

于带约束的目标函数优化方法: LPAm^[10]、LPAm+^[11] 和 LPAh^[9]. 利用目标函数的优化提高社区划分质量, 但没有解决 LPA 稳定性差的问题. 基于节点在网络中的重要性对 LPA 算法进行改进的方法: NIBLPA^[12]、IPLPA^[13] 和 S-LPA^[14]. 根据节点重要性计算标签的影响力, 进而改变算法的更新序列和标签更新方法. 然而, 这些算法没有考虑到网络中节点之间的相似性. 基于节点在网络中的相似性进行改进的方法: NILPA^[15] 和 HPI-LPA^[16]. 通过计算节点与邻居之间的相似性获取节点的更新标签, 提高算法的准确性.

2 LPA_IS 算法

2.1 节点重要性

节点重要性能衡量节点在网络中的影响程度. 主要的方法有: 度中心性、介数中心性、接近中心性、k-shell 算法^[17] 等.

k-shell 算法能够反映出节点在网络中的全局影响力, 但不能反映出同一层节点之间的差异和节点在网络中的局部影响力, 以及节点间的联系程度. 本文基于文献^[13] 和集聚系数^[18] 提出了一种综合衡量节点重要性的方法如式 (2) 所示.

$$NI(i) = \sum_{j \in N(i)} (\alpha_1 \times NKsd(j) + \alpha_2 \times C_j) \quad (2)$$

其中, $NKsd(j)$ 表示 $Ksd(j)$ 进行归一化后的值, C_j 表示节点 j 的集聚系数, α_1 和 α_2 的取值是 0 到 1 之间的常数, $NI(i)$ 表示节点 i 的重要性. Ksd 与 $NKsd$ 的计算公式如下所示:

$$Ksd(i) = Ks(i) + t \quad (3)$$

$$NKsd(i) = \frac{Ksd(i)}{\max\{Ksd(j) | j \in V\}} \quad (4)$$

其中, $Ks(i)$ 表示节点 i 的 ks 值^[17], t 表示删除节点 i 时的迭代次数, $Ksd(i)$ 表示融合节点 i 迭代次数的 ks 值. 集聚系数 C_i 的计算如式 (5) 所示.

$$C_i = \frac{2 \times e_i}{k_i \times (k_i - 1)} \quad (5)$$

其中, e_i 表示节点 i 与其任意两个邻居节点形成的三角形个数, k_i 表示节点 i 的度数. C_i 值越大说明邻居节点之间联系越紧密.

2.2 节点相似性

在标签选择时, 本文将节点相似性作为衡量标签

综合影响力的标准之一. 根据网络的局部信息计算节点相似性的方法有以下两种: 基于共同邻居数的 CN 指标^[19], 基于共同邻居节点度数的指标具体有 Jaccard 指标^[19]、HPI 指标^[20]、RA 指标^[21]等.

本文选取 RA 指标作为计算节点相似性的方法. RA 指标的计算如式 (6) 所示, $s(i, j)$ 表示节点 i 和 j 的相似性, $s(i, j)$ 值越大说明节点 i 和 j 的相似性程度越高, 则它们越有可能属于同一个社区.

$$s(i, j) = \sum_{c \in N(i) \cap N(j)} \frac{1}{k_c} \quad (6)$$

2.3 种子节点集

本文基于节点重要性升序排列获取算法的更新序列, 并根据式 (7) 计算种子节点集, 其中, $i \in V$, V 表示网络的节点集, S 表示种子节点集. 在算法初始化时, 对选取的种子节点分配唯一的标签. 在标签更新过程中, 重要性值小的节点会受到邻居中种子节点的影响, 则这些节点的标签将更新为种子节点的标签, 这有利于提高种子节点在网络中的影响力和减少算法的迭代次数.

$$S = \left\{ i \mid NI(i) > \frac{\sum_{j \in V} NI(j)}{n} \right\} \quad (7)$$

2.4 综合影响力

在标签更新过程中, 节点选择其邻居中综合影响力最大的标签, 作为节点的待更新标签. 本文基于节点重要性与相似性提出了一种计算标签综合影响力的方法. 标签综合影响力的计算如式 (8) 所示.

$$CI(l) = \frac{NI_l(i)}{\sqrt{\sum_{j \in N(i)} NI_l(j)^2}} + \frac{s(i, N_l(i))}{\sqrt{\sum_{j \in N(i)} s(i, j)^2}} \quad (8)$$

其中, $NI_l(i)$ 表示邻居中标签为 l 的节点重要性之和, $N_l(i)$ 表示邻居中标签为 l 的节点集, $s(i, N_l(i))$ 表示节点 i 与其邻居中标签为 l 的相似性之和, $CI(l)$ 表示邻居标签为 l 的综合影响力. 本文使用函数 $u(x) = x / \sqrt{x^2}$ 对 $NI_l(i)$ 和 $s(i, N_l(i))$ 进行归一化处理, 使其能正确反映不同作用力的综合结果.

2.5 算法描述

本文提出的 LPA_IS 算法, 首先根据式 (2) 计算节点重要性, 并获取算法的更新序列. 然后, 基于节点重要性根据式 (7) 计算网络的种子节点集, 并为任意种子节点分配唯一的标签. 最后, 根据式 (8) 计算节点邻居中标签的综合影响力, 通过式 (9) 选择综合影响力最大

的标签作为节点的待更新标签. 节点标签的更新公式如下所示.

$$l_i = \arg \max_{l \in T(i)} CI(l) \quad (9)$$

LPA_IS 算法的具体描述如算法 1 所示.

算法 1. LPA_IS

输入: 网络 $G=(V, E)$, V 表示网络的节点集, E 表示网络的边集, T 表示最大迭代次数
输出: 社区划分结果

1. 通过式 (2) 计算节点重要性, 其中设置 α_1 为 0.45, α_2 为 0.55
2. 按节点重要性升序作为算法的遍历序列 V'
3. 根据式 (7) 计算网络的种子节点集, 并为任意种子节点分配唯一的标签
4. 设置迭代次数 $t=1$
5. while $t < T$ or $can_stop(G) == True$:
6. for each $i \in V'$ do:
7. for $j \in N(i)$ do:
8. 若节点 j 的标签不存在, 则跳过下述操作
9. 计算邻居标签为 l 的节点重要性之和 $NI_l(j)$
10. 计算节点 i 与邻居标签为 l 的相似性之和 $s(i, N_l(i))$
11. end for
12. 根据公式 (8) 计算邻居标签的综合影响力 CI
13. 根据公式 (9) 更新节点 i 的标签
14. end for
15. $t += 1$
16. 如果任意节点标签趋于稳定或达到最大迭代次数, 则停止更新
17. end while
18. 对步骤 5 的结果进行划分, 输出社区划分结果

2.6 复杂度分析

设 n 为网络的节点数, m 为网络的边数, k 为网络的平均度数, S 为网络的种子节点集. ① 计算节点重要性的复杂度为 $O(n)$, 获取种子节点集的复杂度为 $O(n)$, 对任意种子节点初始化标签的复杂度为 $O(|S|)$, S 的大小远远小于 n , 故可忽略不计, 利用桶排序对节点重要性升序排列的复杂度为 $O(n)$, 因此 LPA_IS 算法初始化的复杂度为 $O(3 \times n) = O(n)$; ② 计算节点与其邻居之间相似性的复杂度为 $O(k \times k)$, 利用式 (8) 计算标签综合影响力的复杂度为 $O(k)$; ③ 每一次迭代, 节点进行标签更新的复杂度为 $O(n \times (k \times k + k)) = O(n \times k \times k)$. 假设算法的迭代次数为 T , 故 LPA_IS 算法的总时间复杂度为 $O(T \times k \times k \times n + n)$.

3 实验结果与分析

实验所使用的对比算法包括 LPA、LPAm、LPAh、HPI-LPA 和 S-LPA, 均用 Python 实现, 部分算法存在随机性, 实验结果取运行 100 次后的均值. 实验环境为

Windows 10 操作系统, 内存 8.00 GB, 硬件配置 Inter(R) Core(TM) i7-7770 CPU, 4.00 GB.

3.1 实验数据

实验选取 9 个真实网络数据集, 它们来自不同的真实应用场景^[9,14,22,23], 网络具体信息如表 1 所示.

表 1 真实网络数据集

ID	网络名称	节点数	边数
R1	karate	34	78
R2	dolphins	62	159
R3	polbooks	105	441
R4	football	115	613
R5	power	4941	6594
R6	ca-GrQc	4158	13428
R7	ca-HepTh	8638	24827
R8	PGP	10680	24316
R9	ca-CondMat	21363	91342

本文采用由 Lancichinetti 等人^[24]提出的 LFR 基准程序人工合成 2 组网络, 具体参数设置如表 2 所示, 其中, N 表示节点数, k 表示平均度数, $maxk$ 表示最大度数, $minc$ 表示最小社区内的节点数, $maxc$ 表示最大社区内的节点数, mu 表示社区的混合参数, mu 值越大表示网络结构越模糊, 算法越难识别社区结构.

表 2 人工合成网络数据集

网络名称	N	k	$maxk$	$minc$	$maxc$	mu
N1	5000	15	20	20	80	0.4-0.7
N2	10000	15	30	40	100	0.4-0.7

3.2 评价指标

3.2.1 标准化互信息

若已知网络的真实划分, 则使用标准化互信息 (Normalized Mutual Information, NMI)^[25] 作为评价指标. NMI 的取值为 $[0, 1]$. 若 NMI 值为 1, 则说明算法划分后的社区与网络的真实划分相同. NMI 的定义如式 (10) 所示, 其中, X 表示网络的真实划分, Y 表示算法得到的

社区划分, $H(X|Y)$ 表示 X 在 Y 上的规范化条件熵.

$$NMI(X|Y) = 1 - \frac{H(X|Y) + H(Y|X)}{2} \quad (10)$$

3.2.2 平均模块度

当网络的真实社区划分未知时, 选取模块度 Q ^[26] 作为实验的评价指标. 假设算法运行了 t 次, 每次得到的模块度值为 Q_1, Q_2, \dots, Q_t , 采用 Q_{avg} 表示平均模块度, Q_{avg} 的计算方式如式 (11) 所示.

$$Q_{avg} = \frac{\sum_{i=1}^t Q_i}{t} \quad (11)$$

3.2.3 模块度标准差

在平均模块度 Q_{avg} 的基础上, 本文还采用模块度标准差 Q_{std} 作为实验的评价指标. Q_{std} 的计算方式如式 (12) 所示.

$$Q_{std} = \sqrt{\frac{\sum_{i=1}^t (Q_i - Q_{avg})^2}{t}} \quad (12)$$

3.3 实验过程和结果分析

(1) 真实网络上的实验

从表 3 中可以看出, 在平均模块度的对比上, LPA_IS 算法在 R1、R3 和 R4 网络上的 Q_{avg} 值低于其它算法, 而在其它网络上的 Q_{avg} 值比其它 5 种算法好. 在模块度的标准差方面, LPA_IS 算法的 Q_{std} 值都为 0, 说明本文提出的 LPA_IS 算法具有较好的稳定性. 在真实网络数据集中, R1、R2、R3 和 R4 网络存在真实的社区划分, 在不同算法上比较数据集的 NMI 值, 从表 4 中可以看出, LPA_IS 算法在 R3 网络上的 NMI 值比其它 5 种算法好, 而在 R1 网络上的 NMI 值与 S-LPA 算法相同都为 1, 说明在 R1 网络上的社区划分与真实网络划分一致.

表 3 真实网络模块度对比

ID	LPA		LPAm		LPAh		HPI-LPA		S-LPA		LPA_IS	
	Q_{avg}	Q_{std}	Q_{avg}	Q_{std}	Q_{avg}	Q_{std}	Q_{avg}	Q_{std}	Q_{avg}	Q_{std}	Q_{avg}	Q_{std}
R1	0.3415	0.0209	0.3647	0.0253	0.3872	0.0123	0.3942	0.005	0.3715	0	0.3715	0
R2	0.4827	0.0311	0.4964	0.0300	0.5184	0.0032	0.5238	0.0079	0.4833	0	0.5265	0
R3	0.5043	0.0135	0.5059	0.0126	0.5213	0.0121	0.5141	0.0087	0.4509	0	0.5114	0
R4	0.5884	0.0085	0.5894	0.0128	0.5807	0.0219	0.6010	0.0179	0.6020	0	0.5719	0
R5	0.5941	0.0028	0.5952	0.0039	0.7247	0.0022	0.7223	0.0016	0.7709	0.0005	0.8080	0
R6	0.7174	0.0047	0.7175	0.0034	0.7280	0.0028	0.7585	0.0028	0.7302	0.0001	0.7880	0
R7	0.6042	0.0131	0.6056	0.0060	0.6220	0.0061	0.6489	0.0094	0.4633	0	0.6554	0
R8	0.7328	0.0071	0.7316	0.0092	0.7686	0.0047	0.7905	0.0064	0.7494	0	0.8317	0
R9	0.5697	0.0035	0.5705	0.0032	0.5872	0.0023	0.6078	0.0031	0.5567	0.0001	0.6376	0

在算法运行的平均迭代次数方面,从图1中可以看出,虽然LPA_IS算法在R2、R3和R5网络上的平均迭代次数多于其它算法,但在其它网络上的平均迭代次数比其它5种算法少,这是因为LPA_IS算法利用标签综合影响力更新节点标签,避免了很多不必要的更新。

表4 真实网络 NMI 对比

ID	LPA	LPAm	LPAh	HPI-LPA	S-LPA	LPA_IS
R1	0.6266	0.6786	0.7206	0.6175	1.0	1.0
R2	0.7555	0.7649	0.8384	0.8620	0.7840	0.8495
R3	0.5271	0.5220	0.4859	0.5072	0.5442	0.5544
R4	0.8874	0.8918	0.9150	0.9269	0.9210	0.8138

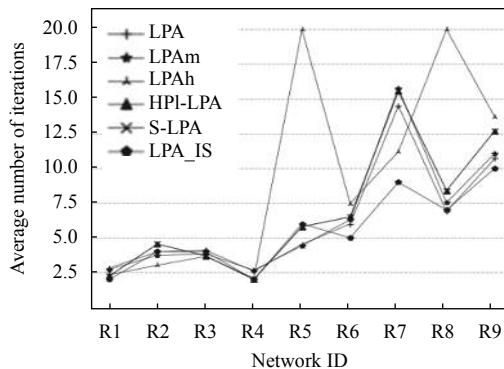


图1 真实网络平均迭代次数对比

综上所述,在真实网络数据集中,LPA_IS算法不仅能够提高算法的准确性和稳定性,而且能够有效地减少算法的迭代次数。

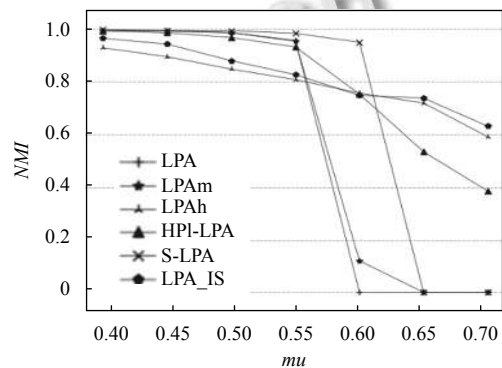
(2) 人工合成网络上的实验

从图2(a)可以看出,当参数 $\mu \leq 0.6$ 时,即社区结构比较明显时,6种算法都能较好的发现社区结构,因此NMI值都比较高。当 $0.6 < \mu \leq 0.7$ 时,即社区结构变得越来越模糊时,LPA_IS算法的NMI值高于其它5种算法。在图2(b)中,当 $0.65 < \mu \leq 0.7$ 时,算法发现社区结构的难度增加,但LPA_IS算法仍能较好地发现社区结构,并且NMI值也都高于其它算法。综上所述,LPA_IS算法相比于其它5种算法,当社区结构变得越来越模糊时,LPA_IS算法仍能保持较高的准确性。

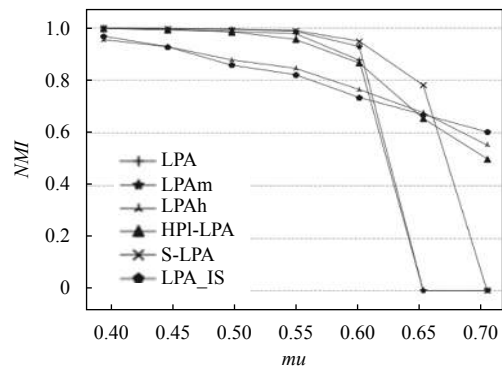
4 总结

针对LPA算法的缺点,本文基于节点重要性与相似性提出了一种改进的标签传播算法(LPA_IS)。算法首先根据改进的k-shell分解算法和集聚系数计算节点

重要性。其次,基于节点重要性提出了种子节点集和算法更新序列的获取方法,并对任意种子节点分配唯一的标签。最后,根据节点重要性与相似性计算标签的综合影响力,任意节点根据其邻居标签的综合影响力更新自身的标签。在真实网络和人工合成网络上的实验结果表明,本文提出的LPA_IS算法与其它5种典型标签传播类算法相比,不仅提高了社区发现的准确性和稳定性,而且减少了算法的迭代次数。



(a) N1 网络上 NMI 对比



(b) N2 网络上 NMI 对比

图2 LFR 基准合成网络上 NMI 对比

参考文献

- 1 Khan MS, Wahab AWA, Herawan T, et al. Virtual community detection through the association between prime nodes in online social networks and its application to ranking algorithms. IEEE Access, 2016, 4: 9614–9624. [doi: 10.1109/ACCESS.2016.2639563]
- 2 Zheng JX, Wang YJ. Personalized recommendations based on sentimental interest community detection. Scientific Programming, 2018, 27(8): 1–14. [doi: 10.1155/2018/8503452]
- 3 Chen YL, Chuang CH, Chiu YT. Community detection based on social interactions in a social network. Journal of the

- Association for Information Science and Technology, 2014, 65(3): 539–550. [doi: [10.1002/asi.22986](https://doi.org/10.1002/asi.22986)]
- 4 Yang JX, Zhang XD. Predicting missing links in complex networks based on common neighbors and distance. *Scientific Reports*, 2016, 6(1): 38208. [doi: [10.1038/srep38208](https://doi.org/10.1038/srep38208)]
 - 5 Girvan M, Newman MEJ. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, 2002, 99(12): 7821–7826. [doi: [10.1073/pnas.122653799](https://doi.org/10.1073/pnas.122653799)]
 - 6 Newman MEJ. Fast algorithm for detecting community structure in networks. *Physical Review E*, 2004, 69(6): 066133. [doi: [10.1103/PhysRevE.69.066133](https://doi.org/10.1103/PhysRevE.69.066133)]
 - 7 Pons P, Latapy M. Computing communities in large networks using random walks. *Proceedings of the 20th International Symposium on Computer and Information Sciences*. Berlin Heidelberg: Springer, 2005. 284–293. [doi: [10.1007/11569596_31](https://doi.org/10.1007/11569596_31)]
 - 8 Raghavan UN, Albert R, Kumara S. Near linear time algorithm to detect community structures in large-scale networks. *Physical Review E*, 2007, 76(3): 036106. [doi: [10.1103/PhysRevE.76.036106](https://doi.org/10.1103/PhysRevE.76.036106)]
 - 9 Luo JH, Ye L. Label propagation method based on bi-objective optimization for ambiguous community detection in large networks. *Scientific Reports*, 2019, 9(1): 9999. [doi: [10.1038/s41598-019-46511-2](https://doi.org/10.1038/s41598-019-46511-2)]
 - 10 Barber MJ, Clark JW. Detecting network communities by propagating labels under constraints. *Physical Review E*, 2009, 80(2): 026129. [doi: [10.1103/PhysRevE.80.026129](https://doi.org/10.1103/PhysRevE.80.026129)]
 - 11 Liu X, Murata T. Advanced modularity-specialized label propagation algorithm for detecting communities in networks. *Physica A: Statistical Mechanics and its Applications*, 2010, 389(7): 1493–1500. [doi: [10.1016/j.physa.2009.12.019](https://doi.org/10.1016/j.physa.2009.12.019)]
 - 12 Xing Y, Meng FR, Zhou Y, *et al.* A node influence based label propagation algorithm for community detection in networks. *The Scientific World Journal*, 2014, 2014: 627581. [doi: [10.1155/2014/627581](https://doi.org/10.1155/2014/627581)]
 - 13 邓凯旋, 陈鸿昶, 黄瑞阳. 基于标签传播能力的改进 LPA 算法. *计算机工程*, 2018, 44(3): 60–64. [doi: [10.3969/j.issn.1000-3428.2018.03.010](https://doi.org/10.3969/j.issn.1000-3428.2018.03.010)]
 - 14 张猛, 李玲娟. 稳定的标签传播社团划分算法研究. *计算机技术与发展*, 2020, 30(1): 129–134. [doi: [10.3969/j.issn.1673-629X.2020.01.023](https://doi.org/10.3969/j.issn.1673-629X.2020.01.023)]
 - 15 Ma TR, Xia ZY. An improved label propagation algorithm based on node importance and random walk for community detection. *Modern Physics Letters B*, 2017, 31(14): 1750162. [doi: [10.1142/S0217984917501627](https://doi.org/10.1142/S0217984917501627)]
 - 16 Song C, Huang GY, Yin B, *et al.* Label propagation algorithm based on node similarity driven by local information. *International Journal of Modern Physics B*, 2019, 33(30): 1950363. [doi: [10.1142/S0217979219503636](https://doi.org/10.1142/S0217979219503636)]
 - 17 Kitsak M, Gallos LK, Havlin S, *et al.* Identification of influential spreaders in complex networks. *Nature Physics*, 2010, 6(11): 888–893. [doi: [10.1038/nphys1746](https://doi.org/10.1038/nphys1746)]
 - 18 Nascimento MCV. Community detection in networks via a spectral heuristic based on the clustering coefficient. *Discrete Applied Mathematics*, 2014, 176: 89–99. [doi: [10.1016/j.dam.2013.09.017](https://doi.org/10.1016/j.dam.2013.09.017)]
 - 19 Aziz F, Gul H, Muhammad I, *et al.* Link prediction using node information on local paths. *Physica A: Statistical Mechanics and its Applications*, 2020, 557: 124980. [doi: [10.1016/j.physa.2020.124980](https://doi.org/10.1016/j.physa.2020.124980)]
 - 20 Ravasz E, Somera AL, Mongru DA, *et al.* Hierarchical organization of modularity in metabolic networks. *Science*, 2002, 297(5586): 1551–1555. [doi: [10.1126/science.1073374](https://doi.org/10.1126/science.1073374)]
 - 21 Zhou T, Lü LY, Zhang YC. Predicting missing links via local information. *The European Physical Journal B*, 2009, 71(4): 623–630. [doi: [10.1140/epjb/e2009-00335-8](https://doi.org/10.1140/epjb/e2009-00335-8)]
 - 22 Leskovec J, Kleinberg J, Faloutsos C. Graph evolution: Densification and shrinking diameters. *ACM Transactions on Knowledge Discovery from Data*, 2007, 1(1): 2–es. [doi: [10.1145/1217299.1217301](https://doi.org/10.1145/1217299.1217301)]
 - 23 Zhang XK, Fei S, Song C, *et al.* Label propagation algorithm based on local cycles for community detection. *International Journal of Modern Physics B*, 2015, 29(5): 1550029. [doi: [10.1142/S0217979215500290](https://doi.org/10.1142/S0217979215500290)]
 - 24 Lancichinetti A, Fortunato S, Radicchi F. Benchmark graphs for testing community detection algorithms. *Physical Review E*, 2008, 78(4): 046110. [doi: [10.1103/PhysRevE.78.046110](https://doi.org/10.1103/PhysRevE.78.046110)]
 - 25 Lancichinetti A, Fortunato S, Kertész J. Detecting the overlapping and hierarchical community structure in complex networks. *New Journal of Physics*, 2009, 11: 033015. [doi: [10.1088/1367-2630/11/3/033015](https://doi.org/10.1088/1367-2630/11/3/033015)]
 - 26 Fortunato S. Community detection in graphs. *Physics Reports*, 2010, 486(3–5): 75–174. [doi: [10.1016/j.physrep.2009.11.002](https://doi.org/10.1016/j.physrep.2009.11.002)]